3900–3907 *Nucleic Acids Research, 1998, Vol. 26, No. 17*

Molecular anatomy of a small chromosome in the green alga *Chlorella vulgaris*

Yoshiteru Noutoshi, Yukari Ito, Shinsaku Kanetani, Makoto Fujie, Shouji Usami and Takashi Yamada*

Department of Molecular Biotechnology, Graduate School of Advanced Sciences of Matter, Hiroshima University, 1-4-1 Kagamiyama, Higashi-Hiroshima 739-8527, Japan

Received June 11, 1998; Revised and Accepted July 20, 1998

DDBJ/EMBL/GenBank accession nos AB013876–AB013884

ABSTRACT

A contig covering the entire region of *Chlorella vulgaris* chromosome I (980 kb long), consisting of 33 cosmid clones has been constructed. By cross-hybridization with other chromosomal DNAs, universal structural elements were detected and localized on the contig. They were composed of at least three different elements: short interspersed DNA elements (SINE)-like elements, long interspersed DNA elements (LINE)-like elements and a putative centromere-like element. At least 36 copies of SINE-like elements were distributed over chromosome I with preferential locations on the right half of the chromosome. DNA fragments containing a SINE-like sequence showed a bent or curved DNA nature on polyacrylamide gel electrophoresis. LINE-like elements were clustered at the left terminus of chromosome I where they formed a tandem array of six copies immediately adjacent to the telomeric repeats. A long sequence element localized at a unique region of chromosome I also existed in a single copy on each chromosome and contained a sequence related to the reverse transcriptase domain of retrotransposons. This feature was compared with the reported centromere-associated elements of higher plants. With its comparative simplicity, the organization of *Chlorella* chromosome I genomic elements may serve as a prototypic experimental system for deciphering the complexity of huge plant chromosomes.

INTRODUCTION

Little is known about the molecular structure and organization of the nuclear genomes of higher plants, although comprehensive investigations are now underway on the sequence composition of plant genomes including the very small (~110 Mb) genome of *Arabidopsis thaliana* (1). Most plant nuclei, however, contain much larger genomes, ranging up to over the 40 000 Mb for *Lilium* species (2). Most of this variation in genome size is attributed to differences in the amounts of repetitive DNA (3). Some of these repetitive sequences are found in tandem satellites, like the chromosomal knobs of maize (4–6), but the majority are interspersed repeats that vary in copy number from tens to thousands per haploid nucleus (7). The nature and organization of these repeats and their functional and structural relationships to genes are not well understood.

Part of what is known about the plant nuclear genome is that each chromosome contains a set of structural components that are essential for its replication, maintenance and segregation. Based on the well established systems of yeast artificial chromosomes (YACs), these chromosomal elements include, at a minimum: replication origins, telomeres and centromeres (8). Some of these elements are shown to have bent or curved DNA structures (9–12).

We are interested in characterizing the molecular organization of fundamental elements in a small plant chromosome. For this purpose, we chose the genome of the unicellular green alga *Chlorella vulgaris*, which is only 38.8 Mb and consists of 16 chromosomes ranging from 980 kb to 4.0 Mb in size (13). Each chromosome can be resolved by pulsed-field gel electrophoresis. The smallest chromosome of this organism (chromosome I, 980 kb in size) can be routinely isolated intact in large quantities. With this chromosomal DNA, we demonstrated that *Chlorella* telomeric repeats are exactly the same as those reported for several higher plants (14). The study of the structural organization of the entire *Chlorella* chromosome should, therefore, provide a unique opportunity to understand what comprises the minimal requirements for plant chromosomes.

We constructed a set of overlapping cosmid clones (contig) for *Chlorella* chromosome I. With these clones we identified and characterized several sequence elements that are conserved in the *Chlorella* chromosomes.

MATERIALS AND METHODS

Strains and growth conditions

Chlorella vulgaris C-169 was obtained from the culture collection of the Institute of Molecular and Cellular Biosciences, University of Tokyo. Cells were cultured photosynthetically in modified Bristol medium (MBM) as described previously (13).

*To whom correspondence should be addressed. Tel: +81 824 24 7752; Fax: +81 824 24 7752; Email: tayamad@ipc.hiroshima-u.ac.jp

Construction of cosmid libraries of *Chlorella* chromosome I

To construct chromosome I-specific cosmid libraries, chromosome I DNA molecules of *C.vulgaris* C-169 were separated by pulsed-field gel electrophoresis as previously described (13). A block of agarose gel containing chromosome I DNA was excised from the CHEF gel and stored in TE buffer at 4°C. For partial digestion with restriction enzymes, two agarose gel blocks (0.25 × 1.0 × 1.5 cm) were equilibrated with restriction buffer for 60 min at 37°C and incubated in 1 ml fresh buffer containing 50 U of *Sau*3AI and incubated at 37°C for 30–60 min. The reaction was stopped by adding 10 µl of 0.5 M EDTA on ice. The agarose gel blocks containing partially digested DNA were then washed with 0.5× TBE buffer. DNA fragments were electro-eluted and ligated to SuperCos1 arms (Stratagene). After packaging with Gigapack III packaging extract (Stratagene), the phages were grown in *Escherichia coli* XL1-Blue MR. The preparation yielded 2×10^4 chromosome I-specific clones.

Alignment of cosmid clones covering the entire region of chromosome I

From the cosmid library, 1200 clones were picked up to generate cosmid contigs and blotted to membranes for hybridization analysis. For the first screening by hybridization, three *Not*I linking DNA clones, the subtelomeric DNA clones from both ends (14) and the gene for α-tubulin (15) were used as landmark probes. The clones were then aligned with each other by *Eco*RI restriction fingerprinting. For further walking, an *Eco*RI fragment positioned far from the first point was used as probe. At each walking step, several independent overlapping clones were obtained.

Detection of sequence elements on chromosome I that are common to all *Chlorella* chromosomes

To detect conserved sequences by Southern hybridization, chromosomal DNA molecules of *Chlorella* cells were separated by pulsed-field gel electrophoresis under two different conditions for smaller and larger chromosomes as described before (13). The chromosomal DNAs separated on the gel were blotted onto a nylon filter (Pall BioSupply) as before (13) and hybridized with labeled individual clones from the cosmid contig of chromosome I as probe. For labeling and hybridization, each cosmid clone DNA was fragmented by digestion with *Eco*RI and *Pst*I. The probes were labeled with fluorescein (Gene Images kit, Amersham) and detected with a CDP-Star detection module (Amersham) according to the manufacturer's protocol. Hybridization was performed at 60°C for 20 h.

For further assignment of sequence elements common to all *Chlorella* chromosomes in a specific restriction fragment of cosmid contig clones, each clone DNA was digested with *Eco*RI and analyzed by Southern blot hybridization. As probe, DNAs of chromosome II, III and VI that were easily isolated by CHEF gel electrophoresis (13) were used. DNAs were digested with *Eco*RI and *Pst*I and labeled with non-radioactive digoxigenin–dUTP by a Boehringer kit according to the manufacturer's manual. Hybridization was performed under a standard condition (14).

Bent DNA isolation and analysis

Bent DNA fragments were screened by two-dimensional polyacrylamide gel electrophoresis according to Mizuno (16). Total *Chlorella* DNA isolated by phenol extraction (14) was digested with restriction enzymes and separated at 60°C in the first dimension and at 4°C in the second dimension. DNA fragments deviated from the gel diagonal were collected from the gel and ligated to appropriate sites of pUC19 for transformation into *E.coli* XL1-Blue MRF′. The bent DNA nature of each clone was confirmed by comparing their electrophoretic mobilities in polyacrylamide gels at different temperatures, as above (16).

DNA sequencing and analysis

Restriction fragments containing chromosome I DNA were cloned into M13 mp18 and 19. Single stranded DNA was sequenced by the chain termination procedure with a kit (Auto Read Sequencing kit, Pharmacia) using an Automated Laser Fluorescence (ALF) DNA sequencer (Pharmacia).

RESULTS

Establishment of a cosmid contig covering the entire region of *Chlorella* chromosome I

A cosmid library of *C.vulgaris* C-169 chromosome I was constructed with DNA isolated by using pulsed-field gel electrophoresis followed by partial digestion with *Sau*3AI. The library consisted of ~20 000 clones with an insert size of 30–42 kb; the combined length of the cloned sequences was 700× the length of chromosome I (980 kb). From a random sample of 1200 clones, a contig map was constructed as follows: the clones were roughly grouped by hybridization with different marker DNA clones that included three *Not*I-linking DNA clones, two subtelomeric clones and the gene clone for α-tubulin whose exact locations were already mapped on chromosome I (14). Each marker assigned an average of 14 cosmid clones. These clones were aligned by comparing their *Eco*RI digestion patterns (*Eco*RI fingerprinting). The technique of 'chromosome walking' identified the gaps between the groups by using overlapping hybridization of clones with partially shared sequences. For each walking step, several independent overlapping cosmid clones were obtained and their relative positions were mapped by restriction analysis. A total of ~400 cosmid clones were definitely aligned along chromosome I, nearly covering it entirely. Any part of the chromosome was covered by several different clones, which helped to rule out the possibility of chimerism or artificial rearrangements in the cosmids. Figure 1 represents the cosmid arrays and overlaps which comprise a minimal tiling (33 clones) of the DNA fragments covering the entire region of chromosome I. Though there were many other clones assigned to the same regions, in the figure, only one representative clone was shown in each region. The most distal parts of the chromosomal ends were covered by the clones previously obtained for the telomeric repeats (pH 2.9 and 1.3) (14). On this cosmid contig, a total of 161 *Eco*RI sites were mapped, so that it is now possible to localize cDNA clones to defined sites on chromosome I by Southern blot. Positions of DNA clones used for landmarks are also shown in Figure 1.

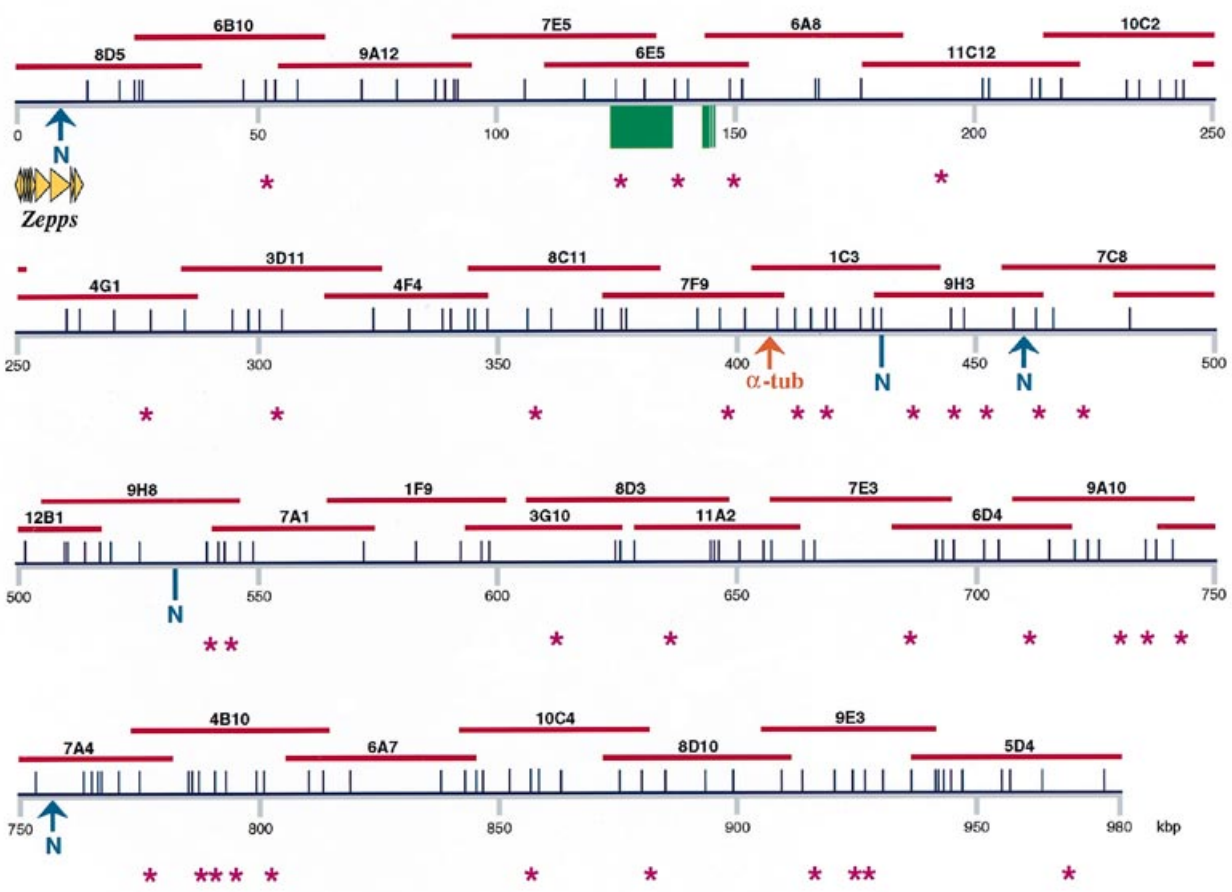

Figure 1. Alignment of minimal overlapping cosmid clones of *Chlorella* chromosome I. The contig consisting of 33 clones is oriented 5′ to 3′ from left to right and constitutive elements, repeats and genes are drawn to scale. The positions of DNA clones used as landmarks including three *Not*I-linking clones (14) and a clone for the α-tubulin gene (15) are shown by arrows. SINE-like elements and LINE-like elements (17) are represented by asterisks and arrowheads (indicated as *Zepp*s), respectively. Boxes in cosmid clone 6E5 show 'common' sequence regions. Restriction sites for *Eco*RI and *Not*I are indicated by upward and downward vertical bars, respectively.

Detection of structural elements on *Chlorella* chromosome I, common to all *Chlorella* chromosomes

When *Chlorella* chromosomal DNAs separated by CHEF gel electrophoresis were hybridized against each of the chromosome I contig clones as probe, several clones from different parts of chromosome I hybridized with almost all of the chromosomes (Fig. 2). For example, clone 8D5 on the left end showed strong hybridization signals with all 16 chromosomes; presumably this was due mainly to the presence of long interspersed DNA element (LINE)-like retrotransposons (*Zepp* elements) in this region (17), because the telomeric repeats were absent from this clone. Six copies of *Zepp* elements formed a tandem array of ~12 kb at the left terminus of chromosome I. A total of 130 copies of *Zepp* elements are known to be distributed over all 16 chromosomes of this organism (18). Similar to the universal elements of *Zepp*, patterns of hybridization with the other *Chlorella* chromosomes and the chromosome I clones, 7E5, 6E5, 6A8, 9H3, 7A1, 7E3, 9A10, 7A4, 4B10, 6A7, 10C4, 8D10 and 9E3 indicated the presence of common sequence elements (hereafter designated 'common sequence elements'). To identify and characterize these elements further, *Eco*RI fragments from each cosmid clone were analyzed using Southern blot hybridization with DNA probes that were fragments from chromosome II, III and IV, which had been isolated from CHEF electrophoretic gels (13). As seen in Figure 3, almost the same hybridization patterns were obtained irrespective of the chromosomes used as probe. The cosmid clones hybridizing with other chromosomes found in Figure 2 consistently gave strong hybridizing bands with all three different probes (Fig. 3). These results implied that some sequence elements common to all *Chlorella* chromosomes are non-randomly distributed on chromosome I and are confined within some *Eco*RI fragments of the cosmid clones.

Assignment of bent or curved DNA sequences on contig clones of chromosome I

We employed an additional strategy to identify chromosomal structural elements on chromosome I. Bent or curved DNA structures are well characterized as components of replication origins (9–11), centromeres (12) and other functional regions (19,20) of chromosomal DNA. Therefore, we attempted to detect and localize such bent DNA elements on *Chlorella* chromosome I. From *Chlorella* DNA fragments separated by two dimensional polyacrylamide gel electrophoresis at different temperatures, we isolated 150 bent DNA clones (21). They were classified into

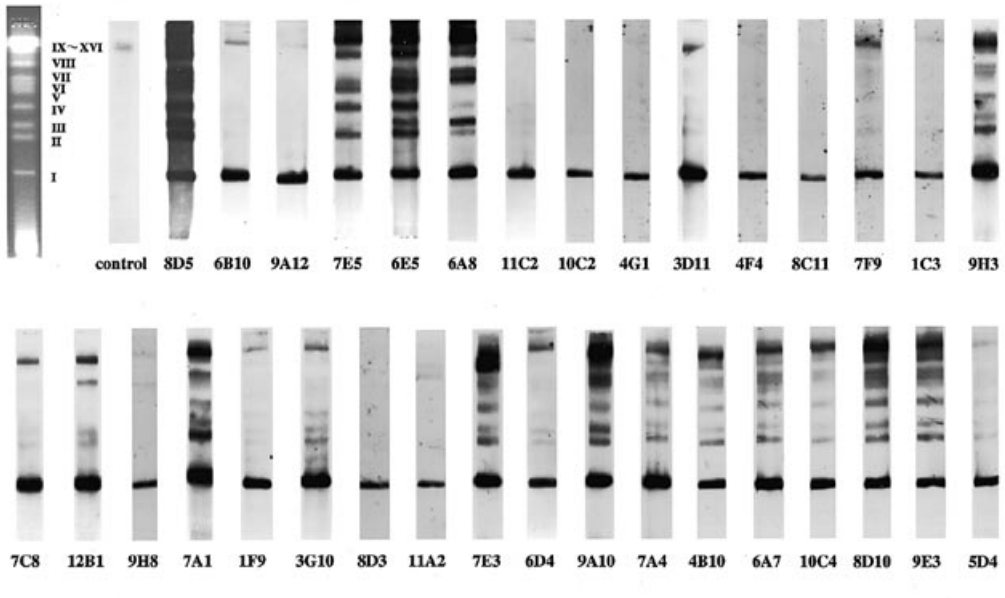

Figure 2. Southern hybridization of *C.vulgaris* chromosomes separated by pulsed-field gel electrophoresis with each clone of chromosome I contig as probe. The chromosome separation pattern is shown in the left-most lane. The clone used as probe is specified under each lane. Cosmid arm DNA was used as a hybridization control probe.

three groups by hybridization with *Chlorella* chromosomal DNAs separated on pulsed-field gels: group I clones hybridized to all the *Chlorella* chromosomes; group II clones hybridized to more than two chromosomes, but not all; group III clones hybridized to only a specific chromosome (21). When clone DH52 belonging to group I was used as a probe for Southern blot analysis of chromosome I contig DNAs, an interesting hybridization pattern appeared. As seen in Figure 4, major hybridizing bands in clones 9H3, 7A1, 7E3, 9A10, 7A4, 4B10, 6A7, 10C4, 8D10 and 9E3 (dotted in Fig. 4B) were exactly the same as those present in Figure 3. These results suggested that the bent DNA elements distributed on chromosome I may be related to some of the common elements detected by chromosomal cross-hybridization, although for clones 7E5, 6E5 and 6A8, the hybridizing bands shown in Figure 4 did not always coincide with the major bands in Figure 3. It is worth noting that these three overlapping clones gave strong hybridization signals with other chromosomes; the signal strength was almost the same for each chromosome (Fig. 2). In these clones, there may be other conserved structural elements that are limited to this region of chromosome I.

Sequence analysis and chromosomal distribution of structural elements present on all chromosomes

To learn about the structural properties of the common elements presented in Figure 3, several bands with strong hybridization signals were excised from the gel, cloned and sequenced. The nucleotide sequences from 3.4 kb of 7C8, 3.4 kb of 7A1, 4.9 kb of 8D10 and 2.8 kb of 9E3 are compared in Figure 5. This figure also includes the sequences of bent DNA clones of group I (21) for comparison. All of these sequences shared an ~200 bp conserved region as indicated in Figure 5. Within this region, some motifs were identified that are characteristic of various short interspersed DNA elements (SINEs), including a pair of flanking repeat sequences, two internally conserved sequence motifs, GATCTG and T/GGG, and a 3′-AT-rich region (22). However, sequences related to the tRNA structures were not obvious at the 5′ region of these elements.

Knowing that the bent DNA clone DH52 was a copy of SINE-like elements that are abundantly distributed over *Chlorella* chromosomes, the results of the Southern blot analysis (Fig. 4) may be interpreted as showing a distribution pattern of such elements on chromosome I. Based on the map of *Eco*RI sites along the cosmid contig clones, the hybridizing bands of Figure 4 could be definitely assigned on chromosome I as shown in Figure 1. Here at least 36 copies of SINE-like elements were localized on chromosome I; they were preferentially distributed on the right half of the chromosome.

Characterization of the common elements confined to cosmid clones 7E5, 6E5 and 6A8

Comparing the results of Figures 2, 3 and 4, we noticed that a region encompassed by cosmid clones 7E5, 6E5 and 6A8 contain other structural elements that are conserved in all *Chlorella* chromosomes. Further analyses by Southern blot revealed that this conserved region was restricted within clone 6E5 and the signals observed with clones 7E5 and 6A8 were due to overlaps with 6E5 (data not shown). To locate the elements precisely, restriction fragments of 6E5 DNA were labeled and used as probe for Southern blot of *Chlorella* chromosomes separated by using pulsed-field gel electrophoresis. The results shown in Figure 6 revealed two separate sub-regions on 6E5 that appeared to be responsible for the conservation; a region spanning consecutive 6.7, 6.0 and 6.5 kb *Eco*RI fragments and another region in an 8.5 kb *Eco*RI fragment. Nucleotide sequences were determined for the ~12.8 kb covered by the 6.7, 6.0 and 6.5 kb *Eco*RI fragments (DDBJ accession no. AB013876). The average base composition of this region was 52.7% GC. A middle part of the region corresponding to restriction fragments 0.9 kb *Hin*dIII, 0.48 kb *Hin*dIII/*Eco*RI and 0.45 kb *Eco*RI/*Xho*I contained several 100–200 bp stretches with considerably high AT contents of >70%. No obviously highly reiterated and clustering sequences

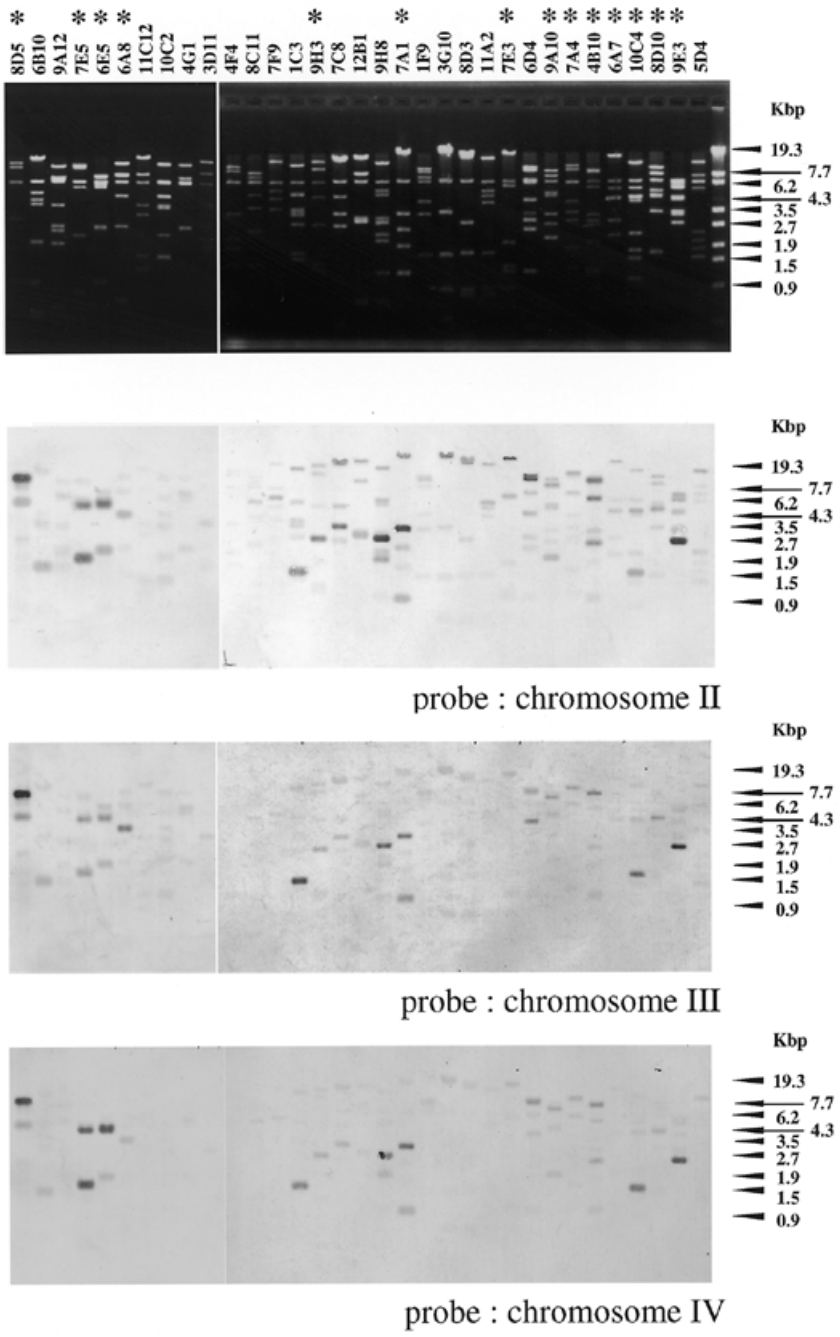

Figure 3. Southern hybridization of *Eco*RI-digested cosmid clones from the chromosome I contig. The blots were hybridized with isolated *Chlorella* chromosomes II, III and IV as probes. The restriction fragmentation patterns are shown in the top part. The same patterns were blotted and hybridized with probes as indicated. Cosmid clones that showed multiple hybridization signals in Figure 2 are indicated by asterisks. Sizes are shown according to *Sty*I-digested λDNA in kilobase pairs.

of more than seven bases, such as telomeric repeats, were present in this region. Although the genes for tRNAs were not present, there was a 141 bp SINE-like element in a 0.91 kb *Apa*I/*Spe*I fragment; cross-hybridization with this fragment was relatively weak, presumably because of the degenerated sequence of this copy.

A total of 11 ORFs in the forward direction and 25 ORFs in the complementary sequence, with a capacity to encode >100 amino acid residues were found in this region. A survey in the databases did not find any significant homologs for these ORFs with a FASTA score higher than 100. The biased codon usage seen in the gene for α-tubulin of this organism (15) suggests that these ORFs are unlikely to encode functional proteins; nevertheless, functional ORFs in this region cannot be excluded. In spite of the absence of prominent features, most restriction fragments derived from this region showed similar hybridization patterns with almost all the *Chlorella* chromosomes (Fig. 6). This suggests the presence in this region of a sequence element that is common to

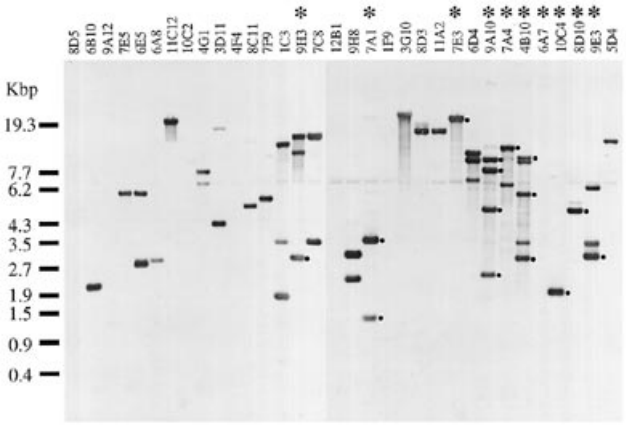

Figure 4. Southern hybridization of *Eco*RI-digested cosmid clones from the chromosome I contig with a bent DNA probe. The blot as in Figure 3 was hybridized with DH52 bent DNA (21) as probe. Clones with asterisks are the same as those in Figure 3. Strongly hybridizing signals also detected in Figure 3 are dotted.

all the *Chlorella* chromosomes, which most likely is in a continuous sequence unit.

Additionally, a central part of the 8.5 kb *Eco*RI fragment of 6E5 covered by the subfragments of a 0.76 kb *Spe*I/*Hin*dIII, a 0.54 kb *Hin*dIII/*Pst*I, a 0.33 kb *Pst*I/*Eco*RV and a 0.7 kb *Eco*RV/*Pst*I was found to contain other conserved sequences (Fig. 6). This region was further divided into a highly conserved left half and a moderately conserved right half judged by hybridization patterns. A total of 2307 bp including these subregions were sequenced (DDBJ accession no. AB013875). The average GC content of this region is 50.6%; regional base compositions are not significantly biased. There are no obviously reiterated and clustering sequence elements, tRNA genes or SINE-like elements within this region. Four ORFs found only in the forward direction showed no significant homology with sequences in the databases except for one ORF (positions 387–962) whose amino acid sequence is homologous with the gene for reverse transcriptase of various retrotransposons (e.g. Rte-1 of *Caenorhabditis elegans*, FASTA score 175). However, this ORF is not closely related to the corresponding gene of *Zepp* elements, a retrotransposon found in *Chlorella*. Because of its truncated size (191 amino acids) compared to functional reverse transcriptase, the expression of this gene is questionable. Since the intensity of hybridization signal with this sequence for individual chromosomes was almost the same as that for chromosome I (Fig. 6), each chromosome should contain only single copy of this sequence. With its location close to the other conserved region, this sequence may have a structural significance in the chromosomal organization.

DISCUSSION

Distribution profiles of SINE-like and LINE-like elements on chromosome I

By chromosome cross-hybridization, we expected to find and localize structural components dispersed on all *Chlorella* chromosomes, such as origins of replication and centromeric elements as well as various kinds of repetitive sequences on *Chlorella* chromosomes. The major common structures detected

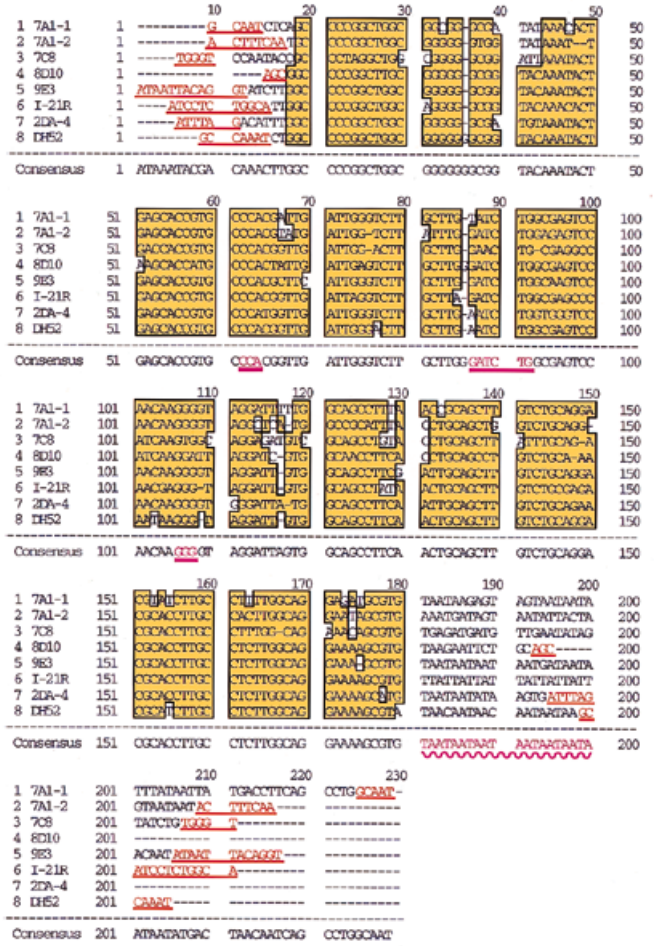

Figure 5. Nucleotide sequence alignment of SINE-like elements on chromosome I and some bent DNA elements. Each element is specified by the name of cosmid clones from which it was derived; 7A1-1 and 7A1-2 are from the same clone 7A1. I-21R, 2DA-4 and DH52 are bent DNA clones belonging to group I (21). Conserved regions are boxed. The motifs reported for typical SINEs (22) are underlined in the consensus sequence. An AT-rich region is indicated by a wavy line. Flanking tandem repeats are underlined for each sequence. The nucleotide sequences have been deposited with the DDBJ database under accession nos AB013877–AB013884.

in this work are of three different types: SINE-like elements, LINE-like elements and some sequences restricted to a specific region. As for SINE-like elements, at least 36 copies are spread over chromosome I. In some locations, they form a cluster of two to four copies; however, they are generally distributed in a single copy, preferentially on the right half of the chromosome. We recently mapped the cDNA clones and found that the right half of chromosome I had a relatively low density of functional genes (Maki *et al*., in preparation). SINE-like elements occupy ~0.7% of the entire chromosome size.

In contrast to the distribution pattern of SINEs, six copies of LINE-like elements (*Zepp* elements) form a tandem array of ~12 kb at the left terminus of chromosome I (17). This *Zepp* cluster is supposed to be a result of repeated integrations of a *Zepp* element into another *Zepp* sequence as a target (17). A similar *Zepp* cluster was also characterized on chromosome V where at least four copies of *Zepp* elements were arranged in a nested

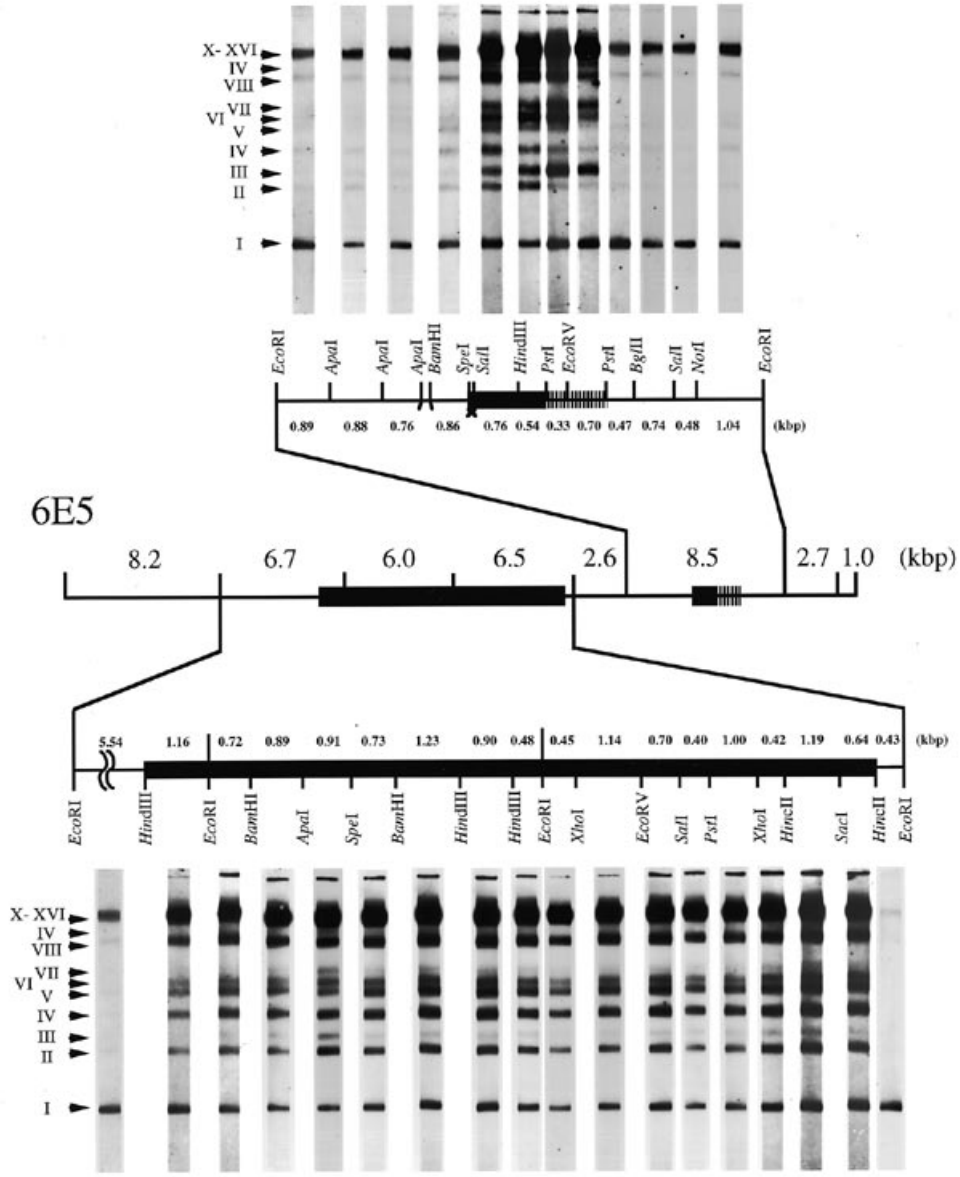

Figure 6. Characterization of common elements on cosmid clone 6E5. Subfragments of 6E5 produced by digestion with several restriction enzymes were hybridized to *Chlorella* chromosomes separated on the pulsed-field gel as probe. The lower part of the figure shows a long conserved region; another conserved region is shown in the upper part. Boxes indicate conserved regions and a vertically striped box shows a moderately conserved region. The positions of each chromosomal DNA on the pulsed-field gel are indicated by Roman numerals.

structure (18). Recently, a full-length copy of this element was isolated and characterized to be 8943 bp in size. Approximately 130 copies of *Zepp*s are estimated to be distributed over the *C.vulgaris* genome (18), corresponding to 1.5% of the total genome size. This value is in accord with the 1.2% *Zepp* occupies of chromosome I.

SINEs, LINEs and a third type of element that was unique to a region of chromosome I, mainly contributed to the cross-hybridization signals in Figure 3. In other words, there were no other significant elements on chromosome I that share common nucleotide sequences in the other chromosomes. Using a standard stringency for hybridization (14), we confirmed that the threshold of detection >80% homology. Several faintly hybridizing signals shown in Figure 3 may reflect additional reiterated sequences of minor quantity. Thus the presence of repeated DNA elements was surprisingly low in the *Chlorella* chromosomes as compared to higher plant genomes. *Chlorella* chromosome I, therefore, may provide a system for studying the minimal requirements for a plant chromosome.

Structural features of SINE-like elements found in chromosome I

SINEs are known as short repetitive elements of ~80–400 bp that are often present in $>10^5$ copies per genome (23). Most of the SINEs so far characterized are from animals but several examples

have also been reported for plants (24). The tobacco SINEs (TS-family) are structurally related to $tRNA^{lys}$ and are considered to belong to the class of tRNA-derived SINEs (24). This type of SINEs is widely distributed in the genomes of dicotyledonous plants including Solanaceae and Convolvulaceae, whereas *Arabidopsis* and several species of monocotyledonous plants do not contain them. The rice SINEs reported by Mochizuki *et al*. (25) are somewhat different from TS-SINEs in their overall structural features; they do not form a composite structure and lack the sequence motifs of GATCTG and TGG. The SINE-like elements found in *Chlorella* chromosomes share common composite structures of tRNA-derived SINEs (22); that is, a pair of flanking tandem repeats, internal conserved motifs of GATCTG and T/GGG, and a 3′-AT-rich region. However, their 5′-regions do not show any significant homology with tRNA structures typical for most SINEs or 7SL RNA sequences for Alu and B1 families (26,27). In these respects, the *Chlorella* elements can be considered as SINE-elements of a novel family. Okada *et al*. (28) proposed as a possible origin of the region homologous to the tRNA of a SINE a primer tRNA, attached to a 'strong-stop DNA', which is an intermediate during reverse transcription of certain retroviruses and long terminal repeat (LTR)-type retrotransposons. They also postulated that an RTase responsible for retrotransposition of SINEs may be provided *in trans* by some LINEs because the 3′ ends of several families of SINEs are homologous with the 3′ ends of some LINEs (29). The structures of *Chlorella* SINE-like elements did not show any significant homology with the LINE elements (*Zepp* elements) found in the same organism. The origin and nature of *Chlorella* SINE-like elements remains to be characterized. The bent DNA nature associated with these elements could be due to sequences rich in AT clusters at the 3′ region of the elements.

The conserved sequence in 6E5

A conserved region found in cosmid clone 6E5 that occurs in other *Chlorella* chromosomes was composed of two sub-domains: a long region of ~12.8 kb and a region containing a reverse transcriptase-like sequence. According to hybridization intensity, this chromosome most likely contains a single copy of both sequence elements. Therefore, it is tempting to speculate that this region might be involved in centromeric function. Nucleotide sequence elements associated with the centromeric regions have been reported for a few plant chromosomes. The 180 bp repeat family, comprising the largest fraction of the highly repetitive DNA (30), is positioned around the centromere on each chromosome of *A.thaliana* (31–33). DNA regions flanking these repeats are reported to be enriched in some retroelement sequences (34). Some other repetitive sequences are also known in regions of the centromeric heterochromatin of *A.thaliana* (35,36). Interestingly, these sequences are related to the reverse transcriptase domain of retrotransposons. A similar case is also reported for cereal plants; DNA sequences homologous to a repetitive DNA element (745 bp long) are located in the centromeric region of sorghum, barley, rye and oats (37). On this 745 bp sequence element, we found an ORF extending over the entire region that has a strong homology with the reverse transcriptase domain of Ty3/gypsy retrotransposons (unpublished data). Although only a few examples are characterized, it is interesting that sequences related to the reverse transcriptase domain of retrotransposons are found in the centromeric regions of chromosomes from different plant species. In this context, the sequence elements of 6E5 found in all *Chlorella* chromosomes are remarkable because of their homology to a reverse transcriptase sequence of retrotransposons. The structural significance of this conserved region might be determined by construction and introduction of an artificial chromosome based on the chromosome I components into *Chlorella* cells and monitoring its stability.

ACKNOWLEDGEMENT

This work was supported in part by a grant from Technical Research Center, Chugoku Electric Co., Inc.

REFERENCES

1 Ecker,J.R. (1998) *Nature*, **391**, 438–439.
2 Friedman,B.E., Bouchard,R.A. and Stern,H. (1982) *Chromosoma*, **87**, 409–424.
3 Flavell,R.B., Bennett,M.D., Smith,J.B. and Smith,D.B. (1974) *Biochem. Genet.*, **12.** 257–269.
4 Peacock,W.J., Dennis,E.S., Rhoades,M.M. and Pryor,A.J. (1981) *Proc. Natl Acad. Sci. USA*, **78**, 4490–4494.
5 Alfenito,M.R. and Birchler,J.A. (1993) *Genetics*, **135**, 589–597.
6 Kaszas,E. and Birchler,J.A. (1996) *EMBO J.*, **15**, 5246–5255.
7 Flavell,R.B. (1980) *Annu. Rev. Plant Physiol.*, **31**, 569–596.
8 Burke,D.T., Carle,G.F. and Olson,M.V. (1987) *Science*, **236**, 806–812.
9 Koepsel,R.R. and Khan,S.A. (1986) *Science*, **233**, 1316–1318.
10 Synder,M., Buchman,A.R. and Davis,R.W. (1986) *Nature*, **324**, 87–89.
11 Bouylikas,T. (1996) *J. Cell Biochem.*, **60**, 297–316.
12 Ueki,N., Momoi,H., Yamada,H. and Mizuno,T. (1993) *Gene*, **132**, 247–250.
13 Higashiyama,T. and Yamada,T. (1991) *Nucleic Acids Res.*, **19**, 6191–6195.
14 Higashiyama,T., Maki,S. and Yamada,T. (1995) *Mol. Gen. Genet.*, **246**, 29–36.
15 Yamada,T., Maki,S. and Higashiyama,T. (1993) *Plant Physiol.*, **103**, 1467.
16 Mizuno,T. (1987) *Nucleic Acids Res.*, **15**, 6827–6841.
17 Higashiyama,T., Noutoshi,Y., Fujie,M. and Yamada,T. (1997) *EMBO J.*, **16**, 3715–3723.
18 Noutoshi,Y., Arai,R., Fujie,M. and Yamada,T. (1998) *Mol. Gen. Genet.* in press.
19 Bouylikas,T. (1995) *Int. Rev. Cytol.*,**162**, 279–388.
20 Traverse,A. (1989) *Nature*, **341**, 184–185.
21 Akiba,M., Noutoshi,Y., Fujie,M. and Yamada,T. (1995) *Nucleic Acids Symp. Ser.*, **34**, 73–74.
22 Okada,N. (1991) *Trends Ecol. Evol.*, **6**, 358–361.
23 Okada,N., Hamada,M., Ogiwara,I. and Ohshima,K. (1997) *Gene*, **205**, 229–243.
24 Yoshioka,Y., Matsumoto,S., Kojima,S., Ohsima,K., Okada,N. and Machida,Y. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 6562–6566.
25 Mochizuki,K., Umeda,M., Ohtsubo,H. and Ohtsubo,E. (1992) *Jap. J. Genet.*, **67**, 155–166.
26 Ullu,E. and Tschudi,C. (1984) *Nature*, **312**, 171–172.
27 Weiner,A.M. (1980) *Cell*, **22**, 209–218.
28 Ohshima,K., Koishi,R., Matsuo,M. and Okada,N. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 6260–6264.
29 Ohshima,K., Hamada,M., Terai,Y. and Okada,N. (1996) *Mol. Cell. Biol.*, **16**, 3756–3764.
30 Martinez-Zapater,J.M., Estelle,M.A. and Somerville,C.R. (1986) *Mol. Gen. Genet.*, **204**, 417–423.
31 Bauwens,S.L., Van Ostveldt,P., Engler,G. and Van Montagu,M. (1991) *Chromosoma*, **101**, 41–48.
32 Maluszynska,J. and Heslop-Harrison,J.S. (1991) *Plant J.*, **1**, 159–166.
33 Murata,M., Ogura,Y. and Motoyoshi,F. (1994) *Jpn. J. Genet.*, **69**, 361–370.
34 Pelissier,T., Tutois,S., Tourmente,S., Deragon,J.M. and Picard,G. (1996) *Genetica*, **97**, 141–151.
35 Brandes,A., Thompson,H., Dean,C. and Heslop-Harrison,J.S. (1997) *Chromosome Res.*, **5**, 238–246.
36 Thompson,R., Schmidt,R., Brandes,A., Heslop-Harrison,J.S. and Dean,C. (1996) *Mol. Gen. Genet.*, **253**, 247–252.
37 Jiang,J., Nasuda,S., Dong,F., Scherrer,C.W., Woo,S.-S., Wing,R.A., Gill,B.S. and Ward.D.C. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 14210–14213.